Specific cognitive signatures of information seeking in controllable environments

Marion Rouault[1,2,3,‡], Aurélien Weiss[1,3,4], Junseok K. Lee[1,3], Jules Bouté[1,3,5], Jan Drugowitsch[6], Valérian Chambon[2,3,*] and Valentin Wyart[1,3,*,‡]

[1]*Laboratoire de Neurosciences Cognitives et Computationnelles, Institut National de la Santé et de la Recherche Médicale (Inserm), Paris, France*

[2]*Institut Jean Nicod, Centre National de la Recherche Scientifique (CNRS), Paris, France*

[3]*Département d'Études Cognitives, École Normale Supérieure, Université Paris Sciences & Lettres (PSL University), Paris, France*

[4]*Université de Paris, Paris, France*

[5]*Paris Saclay Institute of Neuroscience, Paris, France*

[6]*Department of Neurobiology, Harvard Medical School, Boston, MA, USA*

*shared senior authorship

[‡]to whom correspondence should be addressed (marion.rouault@gmail.com, valentin.wyart@inserm.fr)

Number of figures: 5

Number of supplementary figures: 9

Keywords: confidence, changes-of-mind, control, metacognition, volatility, inference.

ORCIDs:

Valérian Chambon: 0000-0002-2401-1906

Jan Drugowitsch: 0000-0002-7846-0408

Marion Rouault: 0000-0001-6586-3788

Valentin Wyart: 0000-0001-6522-7837

Abstract

In uncertain environments, seeking information about the accuracy of alternative strategies is essential for adapting behavior to changes in task contingencies. However, information seeking often co-occurs with changes-of-mind about the perceived accuracy of the current strategy, making it difficult to isolate its specific mechanisms. Here we leveraged the fact that genuine information seeking requires instrumental control to study its cognitive signatures in an adaptive decision-making task tested with and without control. We found that changes-of-mind occurring in controllable environments require more evidence against the current strategy, are associated with reduced confidence, but are nevertheless more likely to be confirmed on the next decision. Computational modelling explained these effects of information seeking through a decrease in the perceived volatility of controllable environments, resulting in stronger and more prolonged effects of changes-of-mind on cognition and behavior. Together, these findings explain the high degree of subjective uncertainty associated with information seeking.

The ability to form and revise uncertain beliefs through information sampling is a hallmark of human cognition under uncertainty. This inference process has been extensively studied using two main classes of decision-making paradigms (Bartolo and Averbeck, 2021; Wyart and Koechlin, 2016). In 'passive sampling' paradigms, participants are observers who sample information over which they have no instrumental control (Murphy et al., 2016; van den Berg et al., 2016; Zylberberg et al., 2018), which is typically the case in perceptual decision-making studies (for reviews, see (Gold and Shadlen, 2007; Hanks and Summerfield, 2017)). In contrast, 'active sampling' paradigms allow participants to choose intentionally which information source to sample (Chambon et al., 2020; Gureckis and Markant, 2012; Markant and Gureckis, 2014; Rouault et al., 2019b). Unlike passive sampling, active sampling provide participants with instrumental control over the information used to form and revise beliefs. As a result, the sampled information corresponds to the outcomes of previous choices made on the basis of the current strategy.

In both types of paradigms, recent work has assigned a key role for subjective confidence evaluation in the flexible formation and revision of ongoing beliefs (Meyniel et al., 2015; Nassar et al., 2010; Rouault et al., 2019a; Sarafyazd and Jazayeri, 2019). Low confidence about current beliefs has been shown to predict changes-of-mind (Balsdon et al., 2020; Folke et al., 2016), and allows participants to flexibly adapt their beliefs and behavior, even when external feedback is unavailable (Desender et al., 2018; Fleming et al., 2018; Rollwage et al., 2020).

Despite this pervasive role of confidence in belief updating, the instrumental control conferred by active sampling produces a fundamental difference between changes-of-mind occurring in controllable and uncontrollable environments. Indeed, active sampling allows to seek information about the accuracy of alternative behavioral strategies, a process that is not possible in the absence of control over information sampling. Consequently, information seeking often co-occurs with changes in behavioral strategy, which makes it difficult to dissociate the mechanisms of pure information seeking from those of changes-of-mind in general.

However, the dependency of information seeking on instrumental control offers experimental leverage to isolate its cognitive signatures, by contrasting changes-of-mind occurring with and without control. Here, we have developed an adaptive decision-making task that compares cognitive inference in active sampling (controllable) and passive sampling (uncontrollable) conditions, while maintaining identical levels of objective uncertainty across conditions (Fig. 1). We found that in controllable environments, participants need more information inconsistent with their current beliefs to change their mind, do so with less confidence, but are also less likely to return to their previous strategy on the next decision. Using computational modelling, we propose that a decrease in the

perceived volatility of controllable environments cause these specific effects of information seeking, resulting in stronger and longer-lasting effects of changes-of-mind on cognition and behavior.

Results

Experimental paradigm comparing passive and active sampling

To examine the specific cognitive signatures of changes-of-mind in controllable environments, we asked human participants to perform an adaptive decision-making task consisting of two tightly matched conditions which only vary in the degree of control over information sampling bestowed to participants (Weiss et al., 2019). Participants were presented with visual sequences of two to eight oriented stimuli drawn from either of two categories (blue or orange) associated with distinct probability distributions over orientation (Fig. 1A). At the end of each sequence, participants were asked to provide a response regarding the category, together with a binary confidence estimate in their response (high or low confidence) using four response keys (see Methods).

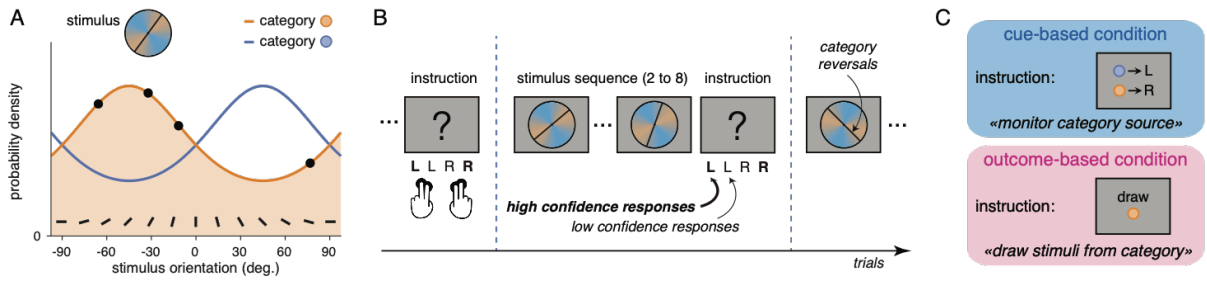

Figure 1. Experimental paradigm probing changes-of-mind across two control conditions

A) Stimuli were bars of orientation drawn from either of two categories (depicted in blue and orange). The generative distributions for each category are depicted. The two category means were orthogonal to each other. Reproduced from Weiss et al., 2019. B) On each trial, participants were presented with sequences of 2 to 8 stimuli drawn from either color category, and were asked to indicate their response and the associated confidence (high, low) using four response keys. Confidence keys (low, high) were assigned to the inner and outer keys of each choice. The target category (hidden state) reversed unpredictably after a pseudo-random number of trials. C) Participants underwent two experimental conditions that varied only in the degree of control over stimuli. In the cue-based condition (blue), observers were asked to monitor the category from which stimuli were drawn. In the outcome-based condition (pink), agents were asked to select an action that will produce stimuli from either category. The evidence available for category choices and visuo-motor properties were tightly matched across conditions (see Methods).

In the passive sampling, cue-based (Cb) condition, participants were asked to monitor the category from which stimuli were drawn, and to report this category at the end of each sequence. The drawn category reversed unpredictably, requiring participants to integrate information across successive sequences to be able to adapt (Fig. 1B). In contrast, in the active sampling, outcome-based (Ob) condition, participants were asked to draw stimuli from either category. Each set of response keys was associated with a target category (e.g., the

left keys with a draw from the blue category and the right keys with a draw from the orange category), and these associations reversed occasionally and unpredictably at the same rate as in both conditions.

In summary, the only difference between conditions is the degree of control participants had over the sampling of stimuli. Participants were observers monitoring the external source of presented stimuli in the Cb condition, whereas they were agents sampling stimuli in the Ob condition (Fig. 1C). Consequently, the hidden state differed between conditions: participants monitored changes in the category being drawn in the Cb condition, whereas they monitored changes in the response key drawing from the target category in the Ob condition.

Psychometric analyses of choice and confidence

To examine differences between conditions, we first analyzed participants' behavior around reversals. After a reversal, participants adapted their behavior to select the response corresponding to the new hidden state (Fig. 2A). We found that participants were slower to adapt in the Ob condition. Indeed, psychometric fits revealed a higher reversal time constant in the Ob than in the Cb condition (t_{32}=5.0, p=1.95e-5) (see Methods). Participants also reached a higher asymptotic reversal rate in the Ob than in the Cb condition (t_{32}=6.3, p=4.0e-7) (Fig. 2B), resulting in slightly higher overall accuracy (correct identification of the hidden state) in this condition (Supplementary Results).

As expected, participants' confidence decreased after a reversal, while participants reacted to the category change (Fig. 2C). Critically confidence decreased more sharply in the Ob relative to the Cb condition (Fig. 2D), as revealed by a difference in a "confidence drop" psychometric parameter (t_{32}=3.59, p=.0011) (see Methods). However, the confidence time constant characterizing the slope of confidence increase after a reversal was only marginally different across conditions (t_{32}=-1.9, p=.064), indicating that confidence rebounded at a comparable speed across conditions (Fig. 2C). Together these findings indicate that in the Ob condition, participants reached a higher plateau level of performance. However, when a reversal occurs, they adapted their choices more slowly and their confidence dropped more significantly than in the Cb condition (Fig. 2).

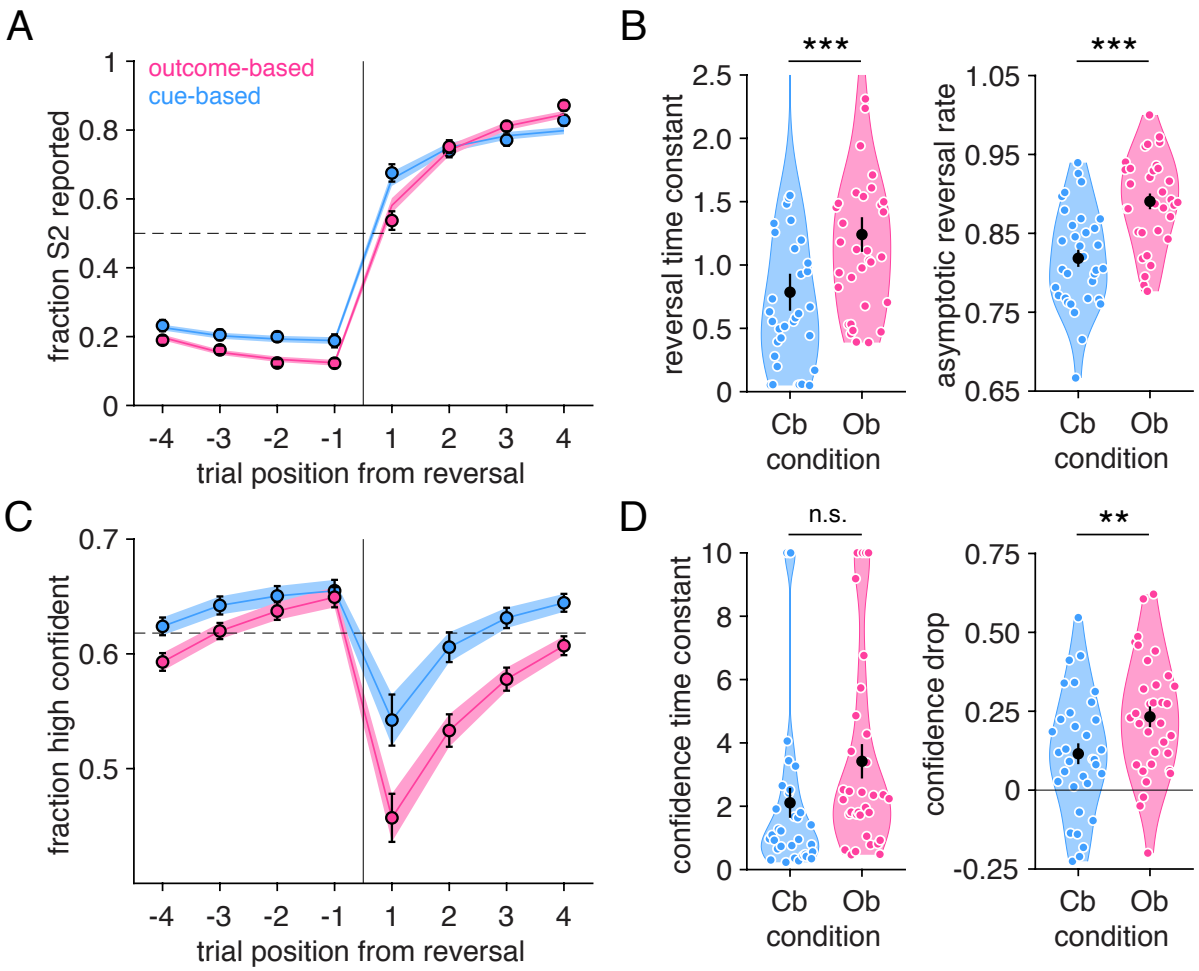

*Figure 2. Behavioral reversal curves characterizing participants' responses and their confidence in cue-based (Cb, blue) and outcome-based (Ob, pink) conditions. Fraction of « S2 » (target hidden state) reported (A) and fraction of high confidence responses (C) as a function of trial number before and after a reversal. Vertical lines indicate the position of reversals. Horizontal dotted lines indicate chance level (A) and mean confidence (C). Circles indicate mean across participants and error bars display S.E.M. across participants (N=33). Shaded areas indicate mean and S.E.M. of psychometric predictions from the best-fitting truncated exponential functions (A) and from the best-fitting sigmoid function (C). (B, D) Psychometric parameters for the fitted response reversal curve (B) and the fitted confidence reversal curve (D). Circles indicate individual participants and error bars indicate the mean and SEM across participants. **p<.01, ***p<.001, n.s. not significant, paired t-tests.*

We then established that participants switched their category choice more often in the Cb (32% of trials) than in the Ob (22% of trials) condition (t_{32}=8.64, p=7.2e-10). To examine why participants were more sticky in the Ob condition, we re-analyzed participants' choice and confidence responses as a function of the objective evidence available in favor of repeating their previous choice ("consistent evidence") vs. switching away from their previous choice ("inconsistent evidence") (Fig. 3). As expected, in both conditions we found that the stronger the consistent evidence, the more often participants repeated their previous choice

(Fig. 3A). Crucially, sigmoid function fits revealed a difference in the point of subjective equivalence (PSE) between conditions (t_{32}=-7.83, p=6.2e-9), a parameter reflecting the amount of evidence required for a participant to shift their category response more frequently than repeat it (Fig. 3B). This PSE difference reveals that participants needed more inconsistent evidence for switching away from their previous choice in the Ob relative to the Cb condition. In contrast, the evidence, whether consistent or inconsistent, was weighted similarly in choices across conditions, as revealed by a lack of a difference in slope between conditions (t_{32}=1.09, p=.28) (Fig. 3B). To sum up, these findings indicate that although participants were equally sensitive to the evidence across conditions, more evidence was required to trigger a switch when participants are in control of the evidence being sampled (Ob condition).

Similarly, we analyzed confidence as a function of consistent vs. inconsistent evidence. As expected, trials with consistent evidence correspond more often to repeat choices (right part of Fig. 3C), whereas trials with inconsistent evidence correspond more often to changes-of-mind (left part of Fig. 3C). We first observed that participants became more confident as the strength of consistent evidence increased (Fig. 3C). In contrast, on choices with inconsistent evidence participants were considerably less confident in the Ob than in the Cb condition (Fig. 3C). To quantify these effects, we fitted two sigmoid functions for switch and repeat choices separately. Each sigmoid was characterized by an offset (PSE) and a slope reflecting sensitivity to the evidence (see Methods). On these two psychometric parameters we performed a 2 × 2 ANOVA with factors Condition (Cb, Ob) and Response type (repeat, switch). We established that on repeat choices, evidence sensitivity was similar across conditions (main effect of Condition, $F(1,32)$=3.8, p=.0596). Confidence was less sensitive to the decision evidence on switch than on repeat trials in both conditions (main effect of Response type, $F(1,32)$=27.9, p=8.6e-6), a decrease in sensitivity that was more pronounced in the Ob condition (interaction between Response type and Condition, $F(1,32)$=8.2, p=.007) (Fig. 3D). For the PSE, we found a main effect of Condition ($F(1,32)$=8.5, p=.006) and Response type ($F(1,32)$=5.3, p=.028), together with an interaction between these factors ($F(1,32)$=7.5, p=.01), revealing that on repeat choices, the PSE was similar across conditions. In contrast, on changes-of-mind the PSE was larger, an increase even more marked in the Ob condition (Fig. 3D).

These findings indicate that in the Ob condition, participants' confidence was similar for repeat trials, when evidence mostly confirmed their prior belief, whereas for switch trials (i.e., mostly inconsistent with their prior belief), participants' confidence dropped as compared to the Cb condition. In addition, we estimated the quantity of evidence corresponding to the smallest fraction of high confidence responses in each condition (akin to the lower point in Fig. 3C). We found that in the Ob condition, participants required more

conflicting evidence for confidence to increase again as compared to the Cb condition (t_{32}=-2.04, p=.05017, jackknifed statistic).

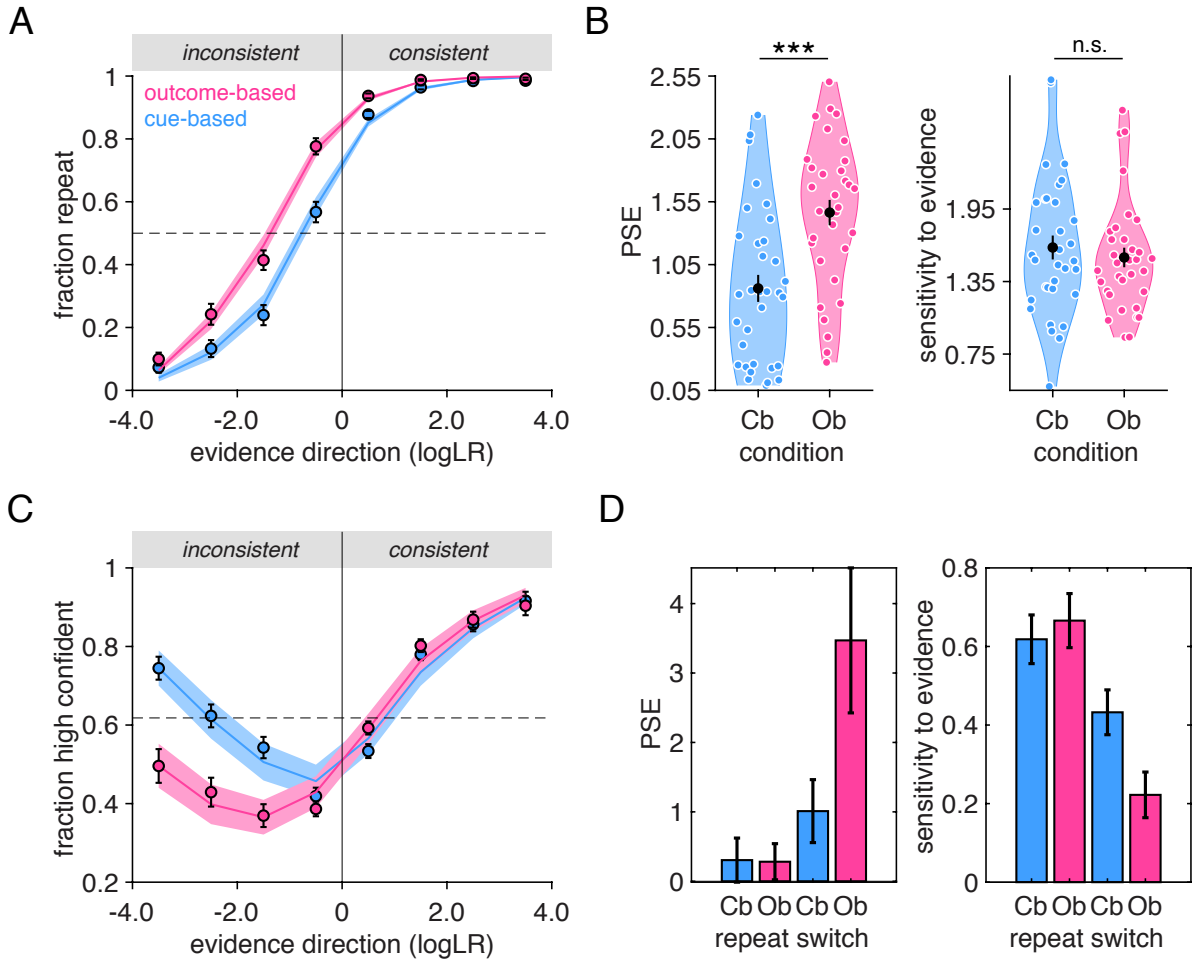

*Figure 3. Behavioral repetition curves characterizing participants' responses and their confidence in cue-based (Cb, blue) and outcome-based (Ob, pink) conditions. Fraction of response repetitions (A) and fraction of high confidence responses (C) as a function of whether the evidence was consistent (in favor of repeating) or inconsistent with the previous choice. Circles indicate human data (N=33) and error bars display S.E.M. across participants (A) or within-subject S.E.M. Within-subject error bars are presented to allow a comparison between conditions without an influence of inter-individual variability about the use and calibration of the high and low confidence responses. (C). Shaded areas indicate mean and S.E.M. of psychometric predictions from the best-fitting truncated exponential functions (A) and from the best-fitting mixture of two sigmoid functions for repeat and switch trials respectively (C). B) Psychometric parameter PSE (point of subjective equivalence) and sensitivity to evidence (slope) parameters for the response repetition curve in B. Circles indicate individual participants and error bars indicate the mean and SEM across participants. ***p<.001, n.s. not significant, paired t-tests. D) Psychometric parameters PSE (point of subjective equivalence) and sensitivity to evidence (slope) for the confidence repetition curves in (C). Bars and error bars indicate jackknifed means and SEM.*

Ruling out alternative accounts of behavioral differences between active and passive sampling

We have seen that participants' confidence dropped more substantially after a reversal in the Ob as compared to the Cb condition (Fig. 2C), and that when changing their mind, participants did so with less confidence than in the Cb condition (Fig. 3C). Although the degree of control over information sampling is the key difference between the two conditions, we sought to examine two alternative interpretations of observed differences between conditions.

First, we examined whether the differences between conditions could be due to a difference in the temporal direction of inference: prospective in the Ob condition (by predicting which key would draw from the target category on the next trial), and retrospective in the Cb condition (by identifying the category of the last sequence of stimuli). We performed a new Experiment 3 to examine this possibility (see Methods). Participants were now asked to guess from which category the computer *drew* from when they saw the sequence of stimuli (retrospective condition), or to guess which category the computer *will draw* from based on what they saw (prospective condition). In other words, neither condition conferred control to participants, but they differed in their temporal direction of inference. We found neither a difference in PSE ($F(1,24)$=3.4, p=.079) nor a difference in confidence drop (t_{24}=.89, p=.38) between conditions (Fig. S2 and Supplementary Results). Together the results provide evidence that the differences between Cb and Ob conditions were not due to a retrospective or prospective temporal orientation, and suggest instead differences due indeed to the degree of control over evidence sampling.

Second, we examined whether the maintenance of a category goal across trials was critical for participants to experience a different degree of control between conditions (Experiment 2B). We created a condition in which the instructions (which category to draw from in the Ob condition, which category/action mapping to use in the Cb condition) changed on a trial basis, instead of on a block basis (see Methods). We found that confidence on trials with inconsistent evidence was less different between conditions (Fig. S8), and the confidence increase after a reversal was similar across conditions (t_{17}=-1.01, p=.33), in line with Experiments 1 and 2B. Moreover, there was a difference in PSE between conditions (t_{17}=-6.4, p=6.4e-6), revealing that participants needed more evidence to change their mind in the Ob condition, again in line with Experiments 1 and 2A (see also Fig. S8 and Supplementary Results). By splitting switch trials into true changes-of-mind about category and simple key press changes without a change in belief about category, we further demonstrate that the drop in confidence for changes-of-mind in the Ob condition was not related to shifts in motor action mapping, but was instead related to true changes-of-mind about category.

A role for confidence and control in changes-of-mind

We have seen that participants changed their mind more often in the Cb than in the Ob condition, indicating a tendency to stick with one's previous choices when participants are in control. Following a change-of-mind, choice and confidence behaviours differed between conditions (Fig. 4A, same data as figure 3C regardless of evidence level). In a 2 × 2 repeated measures ANOVA on the fraction of high confidence responses with Response type (switch, repeat) and Condition (Cb, Ob) as factors, we found that participants were generally less confident when they switched than when they repeated their previous response (main effect of Response type, $F(1,32)$=55.8, p=1.7e-8), and were also less confident in the Ob as compared to the Cb condition (main effect of Condition, $F(1,32)$=17.9, p=.00018) (Fig. 4A). We further identified a significant interaction between Response type and Condition ($F(1,32)$=18.9, p=.00014), indicating a more pronounced decrease in confidence on changes-of-mind in the Ob as compared to the Cb condition. This confidence pattern was mirrored in response times (RTs) (Fig. S3). In a similar ANOVA on RTs, we found that participants were slower on switch than repeat trials (main effect of Response type, $F(1,32)$=65.9, p=2.8e-9) and were also slower when being in control (main effect of Condition, $F(1,32)$=8.27, p=.007). There was an interaction between Response type and Condition ($F(1,32)$=20.0, p=9.1e-5), revealing that a more pronounced slowdown in the Ob condition, in line with a decreased confidence on these trials (Fig. S3).

We focused our next analyses on the consequences of changing one's mind on the following choices. After a change-of-mind, participants could either confirm their category choice on the next trial (the changes-of-mind being "confirmed"), or return back to the category selected before the switch occurred (the changes-of-mind being "aborted"). After changing their mind, participants were less willing to go back to their previous response in the Ob condition, with 72.7% of changes-of-mind confirmed, but only 58.9% in the Cb condition (t_{32}=-8.19, p=2.4e-9). This is consistent with a stronger stickiness tendency observed in the Ob condition (Fig. 4A), suggesting a reluctance to switch back and forth when participants were in control, and instead a willingness to test again their new category choice. Moreover, participants were more confident on trials in which changes-of-mind were confirmed as compared to aborted (2 × 2 repeated measures ANOVA on the fraction of high confidence responses, $F(1,32)$=10.3, p=.0030), and overall more confident in the Cb than in the Ob condition ($F(1,32)$=20.6, p=.0001) (Fig. 4B). We also found an interaction between these factors ($F(1,32)$=10.7, p=.0026), due to participants' confidence decreasing on changes-of-mind aborted as compared to confirmed in the Ob condition. Finally, we observed an equivalent pattern in RTs (Fig. S3). Participants were overall slower in the Ob as compared to the Cb condition (main effect of Condition, $F(1,32)$=6.33, p=.017). Participants were slower to abort than confirm a change-of-mind (main effect of

Confirm/Abort, *F(1,32)*=26.76, *p*=1.2e-5). There was a significant interaction between Condition and Confirm/Abort on RTs (*F(1,32)*=17.2, *p*=.00023), driven by a more pronounced slowdown in aborted changes-of-mind in the Ob condition (Fig. S3). In other words, when not being in control, participants may be more flexible i.e. switch back and forth more easily, as indicated by a similar level of confidence for changes-of-mind confirmed and aborted in the Cb condition (Fig. 4B).

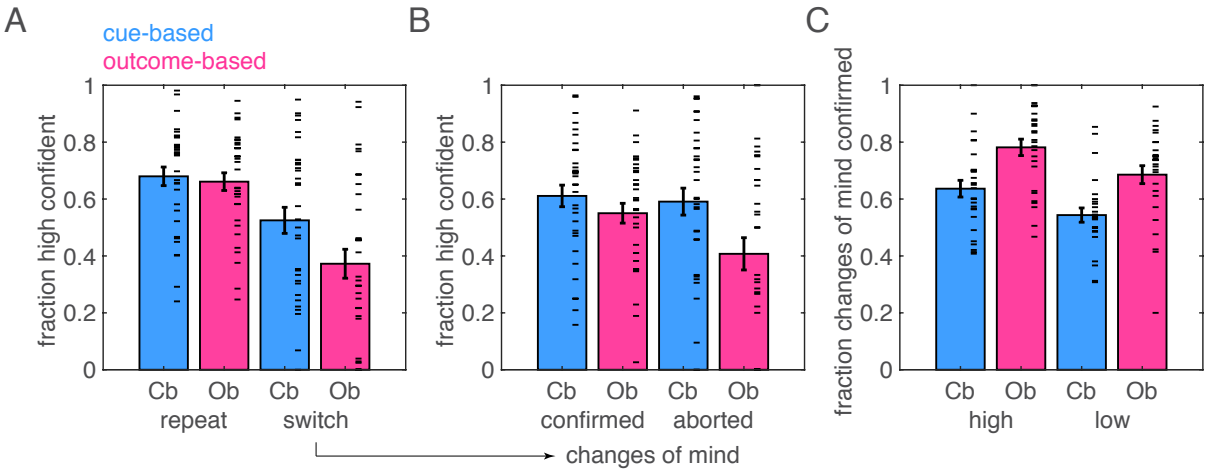

Figure 4. Confidence and changes-of-mind across conditions

A) Fraction of high confidence responses as a function of whether participants repeated their previous choice ("repeat") or changed their mind ("switch") in the cue-based (Cb, blue) and outcome-based (Ob, pink) conditions. B) Fraction of high confidence responses following changes-of-mind that were confirmed or aborted i.e. when participants return back to their previous response in the two conditions. Error bars indicate SEM across participants and black ticks indicate individual data points (N=33). C) Fraction of changes-of-mind confirmed (as compared to aborted, see Methods) as a function of whether the change-of-mind was performed with high or low confidence, for each condition). Bars and error bars indicate mean and SEM across participants (N=26, due to some participants not exhibiting all types of changes-of-mind) and black ticks indicate individual data points.

Similarly, we examined whether changes-of-mind performed with high (resp. low) confidence more often led to choices being confirmed (resp. aborted) in each condition (Fig. 4C). This analysis was restricted to 26 participants due to some participants not exhibiting all types of changes-of-mind. A 2 × 2 repeated measures ANOVA revealed a main effect of Confidence on the fraction of changes-of-mind confirmed (*F(1,26)*=46.2, *p*=3.2e-7), meaning that participants confirmed their switch more often when it was made with high confidence. This suggests a causal role for confidence in controlling changes-of-mind. There was also a main effect of Condition (*F(1,26)*=28.1, *p*=1.5e-5), with changes-of-mind being more often confirmed in the Ob condition, consistent with a higher flexibility of responses in the Cb condition, without an interaction between Confidence and Condition (*F(1,26)*=.069, *p*=.80). Together these findings reveal that participants (i) changed their mind less often in the Ob

condition; (ii) when they did, they did so with reduced confidence; (iii) afterwards, they were more willing to confirm a change-of-mind rather than returning to their previous response, as compared to the Cb condition.

Computational model with inference noise and metacognitive noise

To further characterise the mechanisms underpinning choice and confidence differences between conditions, we developed a normative Bayesian model. In line with previous work on a closely related task, we endowed it with noisy inference (Drugowitsch et al., 2016). In short, the model updates a belief about category from the evidence acquired through the sequence of stimuli, additionally corrupted by inference noise that scales with sequence length, and with a hazard rate controlling reversal occurrence (see Methods). We observed a similar amount of inference noise between conditions (t_{32}=1.45, p=.16), indicating that evidence was integrated equally well across conditions, in line with psychometric results revealing a similar sensitivity to evidence across conditions (Fig. 3). We found a lower perceived hazard rate in the Ob relative to the Cb condition (t_{32}=7.46, p=1.7e-8), indicating that participants perceived the environment as less volatile when being in control (Fig. 5A). To capture the patterns of confidence responses, we further introduced three parameters: a confidence threshold, a metacognitive noise and a confidence gain for switch trials, all capturing different aspects of confidence response patterns (see Methods). Based on the strength of the belief about category, responses above (resp. below) a confidence threshold are given with high (resp. low) confidence, with choices and confidence therefore being based on the same posterior belief. In line with model-free analyses indicating a difference in the fraction of high confidence responses between conditions, we found a lower confidence threshold in the Cb condition, which corresponds to more high confidence responses, as compared to the Ob condition (t_{32}=-4.3, p=1.3e-4), together with no difference in metacognitive noise (t_{32}=-0.29, p=.76), or confidence gain for switch trials (t_{32}=.10, p=.92) (Fig. 5A). Critically, using individual best-fitting parameters, we validated our model by simulating choice and confidence responses, and analysed the simulated data as we did for human data (Palminteri et al., 2017; Wilson and Collins, 2019). We show that simulations provided an excellent fit to participants' choice and confidence responses (Fig. S4). We further validated the independent role of each parameter by median-splitting the participants into groups of high and low hazard rate (Fig. S5), high and low confidence threshold (Fig. S6) and high and low confidence gain for switches (Fig. S7), each of these having a selective influence on participants' choices and confidence. Importantly, even when parameters were similar across conditions (e.g. confidence gain), there was still a substantial inter-individual variability that has a visible effect on participants' confidence (Fig. S7), indicating the

necessity of each parameter in capturing qualitative signatures participants' choice and confidence responses across conditions.

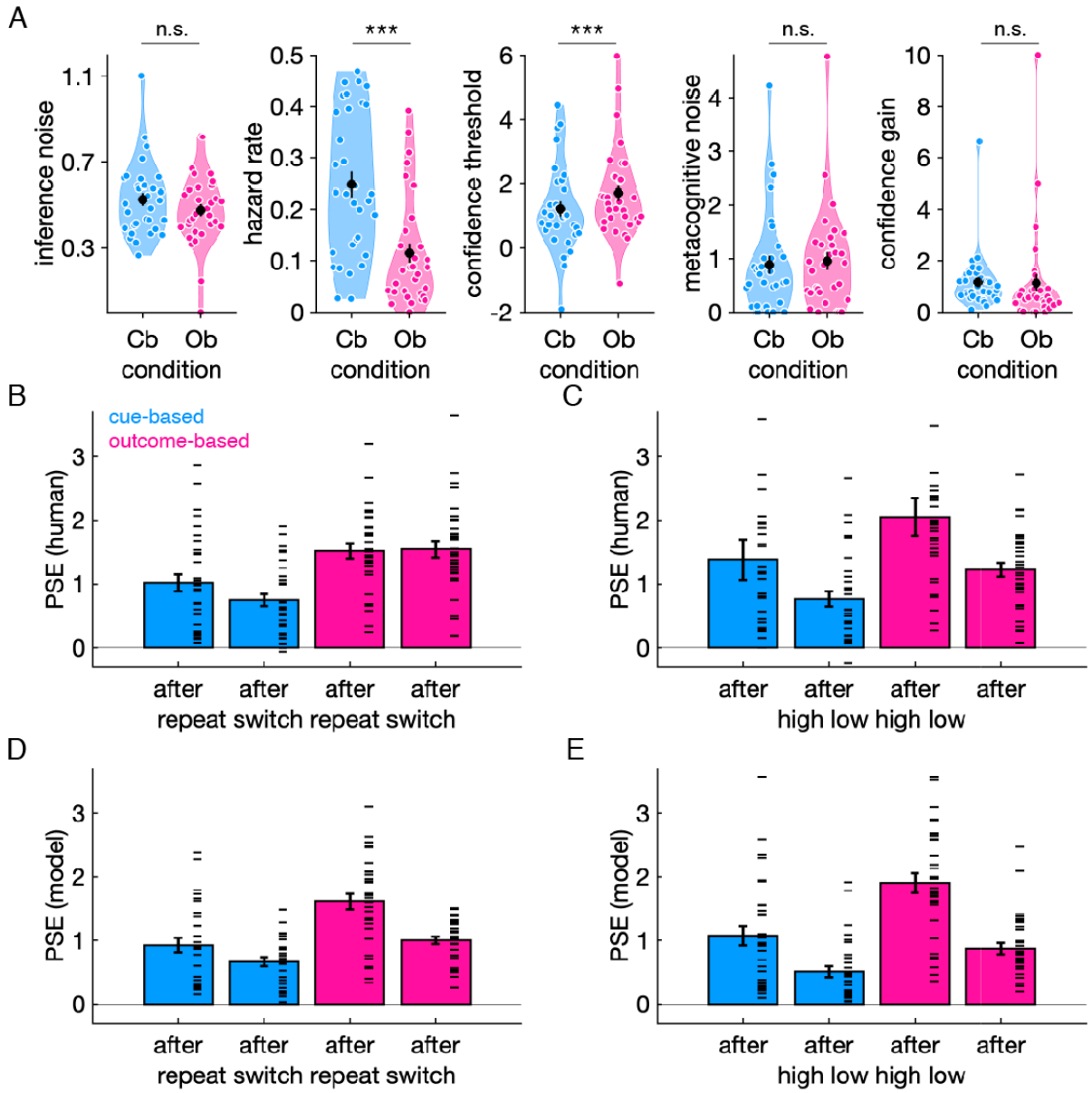

Figure 5. Computational Bayesian model describing choices and confidence

*A) Best-fitting parameters: inference noise, hazard rate, confidence threshold, metacognitive noise and confidence gain for switch trials in cue-based (Cb, blue) and outcome-based (Ob, pink) conditions. Error bars indicate SEM across participants (N=33) and circles indicate individual parameter values (see Methods). ***p<.00001, n.s., non significant, paired t-tests. B-E) Psychometric parameter PSE (point of subjective equivalence) fitted separately on trials following a repeat vs. following a switch, per condition (B, D) and on trials following a high confidence vs a low confidence response, per condition (C, E). Bars and error bars indicate mean and SEM for human data (B, C) or model simulations (D, E) and black ticks indicate individual data points.*

We next examined whether our computational model could not only predict choice and confidence very well (Fig. S4), but also predict post switch effects. Here, we re-fitted the psychometric choice PSE (Fig. 3B) separately for subsamples of trials, for both human data and simulated choice data from the best-fitting set of model parameters for each participant (see Methods). For human choices, in the Cb condition, we found a higher PSE following a repeat as compared to a switch trial, indicating that more evidence was required to switch away from the current best option (t_{32}=3.2, p=.0033), whereas in the Ob condition, PSEs were similar after a repeat and after a switch trial (t_{32}=.08, p=.93) (Fig. 5B). For model choices, in the Cb condition, the PSE pattern was similar to that of human choices (t_{32}=5.4, p=7.0e-6), with a smaller PSE after a switch than after a repeat trial (Fig. 5D). However, for model choices, the PSE was lower in the Ob condition after a switch than after a repeat trial (t_{32}=8.9, p=3.3e-10), unlike for human choices (Fig. 5D). This deviation from model predictions indicates that in controllable environments, participants require more evidence to switch back (following switch trials specifically). This result reveals the origin of a stickiness tendency observed in controllable environments.

We further investigated whether confidence had a role in this deviation. We again fitted PSEs separately for trials following a high or a low confidence response, in each condition (Fig. 5C-5E). As expected, PSEs were lower following a low confidence response, when participants are more willing to revert back to their previous choice. Critically we observed a similar qualitative pattern for human (Fig. 5C) and simulated choices (Fig. 5E). This indicates that the difference in PSE following switches in the Ob condition cannot be due to the model having a reduced metacognitive capacity, because the differences in PSE across confidence levels and between conditions were well captured (Fig. 5C-5E-S4). In sum, we observed a dissociation between changes-of-mind and confidence: while PSE patterns were not different as a function of confidence, a higher PSE was found following a change-of-mind in the Ob as compared to the Cb condition.

Discussion

The ability to adjust confidence in our decisions and change our mind in light of new information is a hallmark of flexible cognition. Previous studies about confidence and changes-of-mind have mostly relied on perceptual decision-making paradigms in which participants are passively sampling information over which they have no control (Fleming et al., 2018; Murphy et al., 2016; van den Berg et al., 2016). However, in many real-life environments, we actively sample information for decision-making, and therefore we experience a degree of control over incoming information (Gureckis and Markant, 2012), a key difference to passive sampling that has not been addressed in previous work. In such controllable environments, information-seeking often co-occurs with changes-of-mind about the reliability of an ongoing strategy. Here, we sought to dissociate genuine information seeking from changes-of-mind. Across several versions of the same experiment, we obtained converging evidence that, in controllable environments, human participants need more evidence to change their mind, change their mind less often and with reduced confidence, and are more willing to confirm a change-of-mind on the next decision than return to their previous strategy. Computational modelling of behavior in controllable and uncontrollable environments indicates that these differences are due to controllable environments being perceived as more stable, resulting in stronger and longer-lasting effects of changes-of-mind on cognition and behavior.

First, we observed that, in controllable environments (outcome-based condition), participants change their mind less often, and need more evidence inconsistent with their previous choice to change their mind. This pattern of effects could indicate that participants change their minds when they have obtained more evidence to do so and should thus end up more confident in their decision to change their strategy. Indeed, previous studies conducted in passive sampling conditions have shown that the stronger the evidence, the higher the confidence (Fleming et al., 2018; Kiani et al., 2014). Instead, in our active sampling condition, we observed that participants are less confident in their changes-of-mind despite, on average, an increased evidence in favor of a change-of-mind in the last sequence of stimuli.

Using computational modelling, we replicate our previous finding of a lower perceived hazard rate in the outcome-based condition (Fig. 5 and S4), corresponding to participants perceived contingencies as more stable in controllable environments (Weiss et al., 2019). Again in line with our previous findings, the decrease in perceived hazard rate was found in the absence of a difference in inference noise between conditions. This pattern indicates that participants process the evidence equally well in both sampling conditions (Drugowitsch et al., 2016; Wyart and Koechlin, 2016). Here we extended the original model to predict subjective confidence reports (Glaze et al., 2015). We followed a normative and

parsimonious approach whereby confidence reports are based on a threshold-dependent readout of the magnitude of the belief used to make category choices. Based on existing work (Maniscalco and Lau, 2012; Pouget et al., 2016), we introduced metacognitive noise in the model as an imperfect readout of the belief. Finally, we accounted for possible asymmetries in the readout of the (log-)belief between choice repetitions and changes-of-mind (see Methods). Through model simulations using the best-fitting parameters of individual participants, we validated that each parameter captures a specific behavioral gradient that matches the psychometric signatures observed in participants (Fig. S4-S7). Importantly, the decrease in perceived hazard rate in the outcome-based condition contributed to explain the reduced confidence for changes-of-mind in this condition. A possible explanation may be that there was an increased conflict between evidence and participants' prior belief about the category choice, hence the confidence decrease in controllable environments. In the model, the stronger prior beliefs triggered by the lower perceived hazard rate require more inconsistent evidence to invert beliefs, likely resulting in lower confidence for changes-of-mind.

The observed effects of controllability on hidden-state inference bear a partial resemblance with the differences between learning under selection vs. reception (Gureckis and Markant, 2012). This earlier study identified a performance benefit for learning under selection where participants can actively sample their environment themselves to test hypotheses about its structure. By doing so, participants maximize the informativeness of their choices, something that they cannot do during learning under reception where participants passively observe yoked choice sequences of another participant in the same task. Similar benefits of active sampling have been reported in terms of memory performance (Voss et al., 2010), and in terms of transfer of prior knowledge to novel object learning (Xu and Tenenbaum, 2007). Active sampling has even been reported to bias perceptual sensitivity in favor of choice-consistent evidence (Talluri et al., 2020). Importantly, in contrast to this earlier work, we designed our paradigm to create tightly matched conditions of active and passive sampling in terms of the information provided, and we offered participants a limited space of hypotheses to choose from (Weiss et al., 2019). This precise manipulation of information presumably explains why we did not observe substantial differences in overall decision accuracy between cue-based and outcome-based conditions. Moreover, the concept of controllable information sampling does not map onto a classical distinction between free and forced choices (Chambon et al., 2020; Sidarus et al., 2019; Wilson et al., 2014). Forced-choice experiments are typically conducted in otherwise controllable environments, and the sequences resulting from free and forced choices therefore vastly differ in terms of the information available for subsequent choices. These differences make it difficult to isolate the specific effects of active sampling over and above its consequences in

terms of decision evidence. In contrast, our paradigm carefully matched the amount of decision evidence provided by cues (passive sampling) and outcomes (active sampling).

In addition, model simulations predicted that the quantity of inconsistent evidence required for participants to switch their response (PSE) should be lower following a response switch than a response repetition, in both conditions (Fig. 5D). Human data systematically violated this prediction in the outcome-based condition in which changes-of-mind are associated with information seeking, with a similar PSE after response repetitions and changes-of-mind (Fig. 5B). This selective deviation between model predictions and human data indicates that participants confirmed their changes-of-mind more often than expected in the outcome-based condition, despite the fact that they were less confident in their changes-of-mind in this condition. These specific behavioral signatures of information seeking could reflect at least two cognitive effects. First, the instrumental control which makes information seeking possible (in the outcome-based condition) may lead participants to perceive a reversal in task contingencies to be less likely following a change-of-mind. Second, participants may engage in a form of hypothesis testing – a specific form of information seeking associated with the sampling of a new source of information – which promotes the repeated selection of the new behavioral strategy after a change-of-mind. Future research should design specific paradigms for arbitrating between these hypotheses.

Using two variants of the same paradigm, we sought to validate controllability as the genuine nature of observed differences between conditions. First, we examined whether the difference in controllability between cue-based and outcome-based conditions could instead be understood as a difference in temporal focus (Experiment 3), with participants performing prospective (future-oriented) inference in the outcome-based condition, and retrospective (past-oriented) inference in the cue-based condition. In “prospective” and “retrospective” conditions carried out in a passive sampling context, we observed that participants do not require more evidence to switch in the prospective condition (Weiss et al., 2019), and that their subjective confidence does not drop more after a reversal in the prospective condition (Fig. S2), two key effects observed in the original outcome-based condition. Together these findings show that differences in temporal focus cannot account for observed differences between cue-based and outcome-based conditions, supporting our interpretation in terms of controllability. Second, we examined whether the maintenance of a fixed category goal throughout a block of trials (e.g., “draw from the orange category”) was critical for participants to perceive controllable environments as more stable. We designed a new task condition in which the instructions (which category to draw from in the outcome-based condition, which mapping between categories and response keys to use in the cue-based condition) changed unpredictably from one trial to the next (Experiment 2B). Despite the change of category goal across successive trials, participants still perceived the outcome-based condition as more

stable than the cue-based condition, suggesting that the behavioral differences observed between conditions arise from instrumental control of information sampling bestowed to participants in the outcome-based condition.

By comparing changes-of-mind occurring in controllable and uncontrollable environments, our findings provide specific cognitive signatures of information seeking that had otherwise proven difficult to isolate. We found that in controllable environments, participants require more evidence to change their mind, do so with reduced confidence but are nevertheless more likely to stick to their subsequent decision. With modelling work indicating participants' perception of controllable environments as more stable, we could not only explain why information seeking is associated with a higher degree of subjective uncertainty, but also identify stronger and longer-lasting effects of changes-of-mind on cognition and behavior. Alterations in confidence (Rouault et al., 2018) and in perceived control (Voss et al., 2017) are a hallmark of various psychiatric symptoms. For instance, participants suffering from schizophrenia present an inflated sense of agency associated with an inability to accurately update these representations (Metcalfe et al., 2014). In contrast, participants suffering from obsessive-compulsive disorder may experience a degraded sense of agency when seeking information in uncertain, controllable environments (Hauser et al., 2017). By clarifying the effects of controllability on inference and confidence, our study lays the groundwork for explaining how the interplay between perceived control and confidence may go awry in psychiatric conditions.

Methods

Participants

Human participants were recruited in the participant pool from the French platform "Relay for Information about Cognitive Sciences" and provided written informed consent. The study was approved by the Comité de Protection des Personnes Ile-de-France VI, ID RCB: 2007-A01125-48, 2017-A01778-45.

Experiment 1. 17 participants were originally recruited in Feb-March 2016. One subject was excluded for aborting the experiment before the end and one for performing at chance level, leaving 15 participants for behavioral analyses.

Experiment 2. 20 participants were initially recruited in Nov-Dec 2016 and in March-May 2017. Two participants were excluded for performing at chance level, leaving 18 participants for behavioral analyses.

Experiment 3. 30 participants were initially recruited in March 2019. Four participants were excluded for performing at chance level and one subject aborted the experiment before the end, leaving 25 participants for behavioral analyses.

Experiment 4. 24 participants were tested in Sept-Nov 2015, whose behaviour is described in (Weiss et al., 2019).

Experiment 5. 34 participants were initially recruited in January-March 2018. No confidence reports were recorded in this data set. Three participants were excluded for performing at chance level in some of the blocks, leaving of total of 31 participants for basic analyses. Three additional participants were excluded from psychometric PSE analyses for not exhibiting all types of events for the analysis (e.g. no switch trials in either condition in volatile blocks), leaving a total of 28 participants for psychometric analyses.

Behavioral tasks

Experiment 1. Participants performed a reversal learning task similar to our previous study (Weiss et al., 2019). Participants were presented with sequences of stimuli drawn from two discrete colour categories (Fig. 1A), and were asked to make a decision about the generating category on each trial (each sequence) (Fig. 1B). To examine a role for subjective confidence in relation to inference and changes-of-mind about category, in addition to their

choice, participants indicated their confidence (high or low) in their response using four response keys (Fig. 1B).

Participants performed two experimental conditions that aimed at examining the influence of the degree of control over stimuli on choice and confidence. In the cue-based (Cb) condition, participants were instructed that the computer draws sequences of stimuli, and were asked to identify the category from which the stimuli were drawn. An instruction screen indicated the mapping between response keys and colour categories (counterbalanced across blocks). In the outcome-based (Ob) condition, participants were instructed to draw stimuli from a given category. An instruction screen indicated the target colour category for each block (counterbalanced across blocks). The conditions were otherwise fully symmetric, tightly matched in terms of visual and motor requirements, and the order of condition administration was counterbalanced across participants.

The generating category (hidden state) reversed from time to time, with "episodes" i.e. chunks of trials during which the hidden-state was fixed. Episode duration was sampled pseudo-randomly from a truncated exponential probability distribution (between 4 and 24 trials), resulting in a near-constant hazard rate in each block. Participants completed a total of 576 trials divided into blocks of 72 trials (72 sequences). The detailed instructions provided to participants are reported in (Weiss et al., 2019).

Experiment 2. To examine whether the maintenance of a category goal across longer time-scales was critical in participants' experiencing of control, and therefore influenced the findings, we introduced a new rule manipulation. In half of the blocks, rules were stable across a block (thereafter, Experiment 2A). Since behavior was virtually identical in this rule condition, we then pooled these data with Experiment 1 for a total of 33 participants. In the other half of blocks, rules were changing on a trial basis instead of a block basis (thereafter, Experiment 2B). The generative structure of the task and all other experimental features remained identical to Experiment 1. We analysed this experimental data (2B) separately (Fig. S8).

Experiment 3. We reasoned that our Cb and Ob conditions differ in terms of the degree of control that participants experience over stimuli. However, an alternative interpretation of these differences would be in terms of temporal orientation, with the Cb condition being about monitoring past stimuli retrospectively (instruction to "monitor"), whereas the Ob condition would be about producing stimuli prospectively (instruction to "draw"). We set out to test this hypothesis using a modified version of the original conditions. Here, after each sequence of stimuli, participants were asked to guess from which category the computer drew from (retrospective condition), or to guess which category the computer will draw from

(prospective condition). All other experimental features remained identical to Experiments 1 and 2A.

Experiment 4. Detailed behavioral, modelling and MEG analyses of this experimental data have been presented elsewhere (Weiss et al., 2019). In short, the experimental design was similar to that of Experiment 1 with the exception that no confidence responses were asked, only category choices.

Experiment 5. Since a lower perceived volatility of the environment in the Ob condition was sufficient to explain choice differences across conditions (Weiss et al., 2019), we sought to establish whether these differences would replicate across varying volatility conditions. Here, we experimentally manipulated the hazard rate with half of the blocks presenting a low volatility (true hazard rate = 1/12) and half a high volatility (true hazard rate = 1/6). No confidence estimates were required here, only category choices. Critically, participants were explicitly instructed about volatility conditions, and were cued about volatility level on a block basis (e.g. “Frequent changes”). This manipulation allowed us to test whether participants adapted their point of subjective equivalence (PSE) as a function of volatility instructions (Supplementary Results and Fig. S9).

Stimuli. Stimuli were oriented bars presented on top of a colored circle displaying an angular gradient between orange and blue colour categories spaced by π/2 (Fig. 1A). Stimuli were drawn from either of these two categories (Fig. 1A). On each trial, a sequence of 2, 4, 6 or 8 stimuli was drawn from a von Mises probability distribution centered on either category with a concentration of 0.5, and presented to participants for making a decision about category of origin at the end of the sequence (Fig. 1B). The number of stimuli per sequence was drawn from a uniform distribution and pseudo-randomised across trials. Stimuli were displayed at an average rate of 2 Hz with an inter-stimulus interval of 500 ± 50 ms. The last stimulus of each sequence was followed by a longer delay of 1000 ± 50 ms, before participants were probed for their response.

Statistical analyses and quantitative analyses

Model-free behavioral analyses. We defined performance as the proportion of choices of the true hidden state (true generative color category). We examined whether mean performance and mean confidence (fraction of high confidence responses) varied across Cb and Ob conditions using paired *t*-tests. To assess how much insight participants had into their performance, we computed discrimination ability for each condition as the difference in confidence on correct and error responses (Fig. S1). We also computed calibration, another

measure of metacognitive ability, that reflects overall how closely participants' confidence matched their performance (mean performance minus mean fraction of high confidence responses) (Fig. S1).

Psychometric analysis of behavior. We first examined two reversal curves: the proportion of choosing either option (Fig. 2A) and the proportion of high confidence responses (Fig. 2C) as a function of trial number before and after a reversal. We modelled choices using a truncated exponential function characterised by two parameters, an asymptotic level and a (reversal) time constant (Fig. 2B). For confidence as a function of reversal, we also an exponential learning model for which we report two characteristic parameters (Fig. 2D), a (confidence) time constant, and a "confidence drop" parameter reflecting the drop between the lower (pmin) and upper (pmax) asymptotic levels (corrected by the time constant), such that:

$$Confidence\ drop = pmax - pmin + (pmax - pmin) \times (1 - e^{-1/time\ constant})$$

We also examined the proportion of repeating the previous choice (Fig. 3A) as a function of evidence recoded in favor of repeating a previous choice ("consistent evidence") or in favor of switching choice ("inconsistent evidence"). We quantified this by fitting a logistic function to quantify the amount of evidence required to switch a response in each condition (PSE, point of subjective equivalence) and the sensitivity to the evidence (slope of the sigmoid function) (Fig. 3B). Similarly, we analysed the proportion of high confidence responses as a function of consistent vs. inconsistent evidence (Fig. 3C). For confidence we computed within-subject error bars by removing the mean confidence level across conditions before computing the standard error. This was done to allow a comparison between conditions without an influence of inter-individual variability about the use and calibration of the high and low confidence responses. To quantify these findings we fitted two logistic sigmoid functions, one for repeat and one for switch choices, for each condition separately. We also quantified the evidence level at which the two sigmoid for repeat and switch trials intersect, which corresponds to the quantity of evidence required for confidence to increase again when choices are made with inconsistent evidence (Fig. 3D). For all psychometric analyses we compared parameters of best-fitting functions across conditions using paired *t*-tests (Figs. 2B-2D-3B-3D). When individual estimates were too noisy (e.g. for 6/33 participants, confidence was very little modulated by evidence level on switch trials), we resorted to a jackknifing procedure, akin leave-one-out analyses (Kiesel et al., 2008).

We additionally fitted the PSE on subsamples of trials corresponding to after a repeat and after a switch trials, and trials with high or low confidence. This allowed us to examine how participants adapted their PSE on a dynamic basis, depending on the choice sequence experienced in the task.

Change-of-mind analyses. In Experiments 1 and 2A, we sought to characterise the behavioral properties of changes-of-mind across conditions (Fig. 5). In a 2 × 2 repeated measures ANOVA, we examined the influence of Response type (repeat, switch) and Condition (Cb, Ob) on the fraction of high confidence responses (Fig. 5A). For switch trials, we further examined confidence on switch trials that were confirmed on the next trial ("change-of-mind confirmed") as compared to switch trials after which participants went back to their previous response ("change-of-mind aborted") (Fig. 5B). Finally, we examined the fraction of changes-of-mind confirmed (over all changes-of-mind) as a function of whether the change was performed with high or low confidence (Fig. 5C).

In Experiment 2B, we could further dissociate true changes-of-mind about category, as compared to motor changes-of-mind related to key presses, without a change in belief about category. We report the results for two 2 × 2 repeated measures ANOVAs, one on true changes-of-mind and one on button changes, with Response type (repeat, switch) and Condition (Cb, Ob) as within-subject factors (Fig. S8E).

Computational model

Model structure. We implemented a normative model of perceptual inference in volatile environments with contingency reversals (Glaze et al., 2015). A key aspect of the model is that it proceeds on the same evidence quantities similarly for the Cb and Ob conditions. Since previous work indicates that inference noise and not selection noise explains most of choice variability in such an inference task (Drugowitsch et al., 2016), we included no selection noise but we introduced inference noise on each sample of the sequence, that scales with sequence length (Weiss et al., 2019). We extended the model in important ways to predict not only choices, but also confidence. First, we introduced a confidence threshold parameter for determining whether the response will be provided with high or low confidence based on the posterior belief about the chosen category. This parameter captured baseline differences in proportion of high confidence responses between conditions, representing the most parsimonious extension possible to provide a normative confidence response. Second, we introduced metacognitive noise to model an imperfect readout, as previously proposed (Maniscalco and Lau, 2012; Pouget et al., 2016). Third, we introduced a confidence gain parameter on switch trials modelling a differential readout of the posterior belief on changes-of-mind only, as a multiplicative factor on the posterior belief applied on switch trials only.

Model fitting. We fitted the model using Bayesian Adaptive Direct Search (BADS) with 100 validation samples which provides point estimates for each parameter (Acerbi and Ma, 2017). We used five random starting points and parameters were bounded as follows (hazard rate: range=0.000001-0.999999, inference noise: range=0-10, metacognitive noise:

range=0-10, confidence threshold: range=-10,+10, confidence gain for switches: range=0-10). We maximised the likelihood that model choices and confidence reproduce the four reversal (Fig. 2) and repetition (Fig. 3) curves of participants, that are the important dimensions of interest for understanding participants' choice patterns. One participant was excluded for having an unreliable fit. All parameters were allowed to vary between conditions, and we compared the best-fitting parameters using paired *t*-tests.

Model validation. To validate the model, we simulated model choice and confidence from the best-fitting parameters on the same stimuli sequences as participants, and analysed the simulations similarly as for participants' data (Fig. S4) (Palminteri et al., 2017; Wilson and Collins, 2019). We also performed a median split across participants on the best-fitting hazard rate (Fig. S5), confidence threshold (Fig. S6) and confidence gain (Fig. S7), and averaged simulations of the model for each subgroup, to further illustrate the independent contribution of each of these parameters.

Finally, to compare the characteristics of changes-of-mind between model and participants, we reproduced one of the psychometric analyses that is fitting the choice PSE separately for trials following a repeat vs. a change-of-mind, and separately for trials following a high vs. a low confidence response. For model simulations, we averaged across 50 simulations per participant (Fig. 5).

Data and code availability

All data were collected under ethics agreements set up before GDPR onset, so that participants did not provide consent for their data to be shared on a public repository. However, data can be shared with other researchers upon request as consented by participants. MATLAB code for behavioral analyses and computational modelling will be made available at https://github/marionrouault/actobscom/ upon publication in a peer-reviewed journal.

References

Acerbi, L., Ma, W.J., 2017. Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. Advances in neural information processing systems 30, 1836–1846.

Balsdon, T., Wyart, V., Mamassian, P., 2020. Confidence controls perceptual evidence accumulation. Nature Communications 1–11. doi:10.1038/s41467-020-15561-w

Bartolo, R., Averbeck, B.B., 2021. Inference as a fundamental process in behavior. Current Opinion in Behavioral Sciences 38, 8–13.

Chambon, V., Thero, H., Vidal, M., Vandendriessche, H., Haggard, P., Palminteri, S., 2020. Information about action outcomes differentially affects learning from self-determined versus imposed choices. Nat. hum. behav. 4, 1067–1079.

Desender, K., Boldt, A., Yeung, N., 2018. Subjective Confidence Predicts Information Seeking in Decision Making. Psychological Science 29, 761–778.

Drugowitsch, J., Wyart, V., Devauchelle, A.-D., Koechlin, E., 2016. Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. Neuron 92, 1398–1411. doi:10.1016/j.neuron.2016.11.005

Fleming, S.M., Putten, E.J., Daw, N.D., 2018. Neural mediators of changes of mind about perceptual decisions. Nat Neurosci 1.

Folke, T., Jacobsen, C., Fleming, S.M., De Martino, B., 2016. Explicit representation of confidence informs future value-based decisions. Nat. hum. behav. 1, 105–8. doi:10.1038/s41562-016-0002

Glaze, C.M., Kable, J.W., Gold, J.I., 2015. Normative evidence accumulation in unpredictable environments. eLife 4, e08825.

Gold, J.I., Shadlen, M.N., 2007. The neural basis of decision making. Annu. Rev. Neurosci. 30.

Gureckis, T.M., Markant, D.B., 2012. Self-directed learning: A cognitive and computational perspective. Perspectives on Psychological Science 7, 464–481.

Hanks, T.D., Summerfield, C., 2017. Perceptual decision making in rodents, monkeys, and humans. Neuron 93, 15–31.

Hauser, T.U., Moutoussis, M., Iannaccone, R., Brem, S., Walitza, S., Drechsler, R., Dayan, P., Dolan, R.J., 2017. Increased decision thresholds enhance information gathering performance in juvenile Obsessive-Compulsive Disorder (OCD). PLoS Comput Biol 13, e1005440.

Kiani, R., Corthell, L., Shadlen, M.N., 2014. Choice Certainty Is Informed by Both Evidence and Decision Time. Neuron 84, 1329–1342. doi:10.1016/j.neuron.2014.12.015

Kiesel, A., Miller, J., Jolicœur, P., Brisson, B., 2008. Measurement of ERP latency differences: A comparison of single-participant and jackknife-based scoring methods. Psychophysiology 45, 250–274.

Maniscalco, B., Lau, H., 2012. A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. Consciousness and Cognition 21, 422–430. doi:10.1016/j.concog.2011.09.021

Markant, D.B., Gureckis, T.M., 2014. Is it better to select or to receive? Learning via active and passive hypothesis testing. Journal of Experimental Psychology: General 143, 94–122. doi:10.1037/a0032108

Metcalfe, J., Van Snellenberg, J.X., DeRosse, P., Balsam, P., Malhotra, A.K., 2014. Judgments of agency in schizophrenia: an impairment in autonoetic metacognition, in: The Cognitive Neuroscience of Metacognition. Springer, pp. 367–387.

Meyniel, F., Schlunegger, D., Dehaene, S., 2015. The sense of confidence during probabilistic learning: A normative account. PLoS Comput Biol 11, e1004305.

Murphy, P.R., Robertson, I.H., Harty, S., OConnell, R.G., 2016. Neural evidence accumulation persists after choice to inform metacognitive judgments. eLife 1–23. doi:10.7554/eLife.11946.001

Nassar, M.R., Wilson, R.C., Heasly, B., Gold, J.I., 2010. An Approximately Bayesian Delta-Rule Model Explains the Dynamics of Belief Updating in a Changing Environment. Journal of Neuroscience 30, 12366–12378. doi:10.1523/JNEUROSCI.0822-10.2010

Palminteri, S., Wyart, V., Koechlin, E., 2017. The Importance of Falsification in Computational Cognitive Modeling. Trends in Cognitive Sciences.

Pouget, A., Drugowitsch, J., Kepecs, A., 2016. Confidence and certainty: distinct probabilistic quantities for different goals. Nat Neurosci 19, 366–374. doi:10.1038/nn.4240

Rollwage, M., Loosen, A., Hauser, T.U., Moran, R., Dolan, R.J., Fleming, S.M., 2020. Confidence drives a neural confirmation bias. Nature Communications 11, 1–11.

Rouault, M., Dayan, P., Fleming, S.M., 2019a. Forming global estimates of self-performance from local confidence. Nature Communications 10, 1141.

Rouault, M., Drugowitsch, J., Koechlin, E., 2019b. Prefrontal mechanisms combining rewards and beliefs in human decision-making. Nature Communications 1–16. doi:10.1038/s41467-018-08121-w

Rouault, M., Seow, T., Gillan, C.M., Fleming, S.M., 2018. Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. Biological Psychiatry.

Sarafyazd, M., Jazayeri, M., 2019. Hierarchical reasoning by neural circuits in the frontal cortex. Science 364.

Sidarus, N., Palminteri, S., Chambon, V., 2019. Cost-benefit trade-offs in decision-making

and learning. PLoS Comput Biol 15, e1007326.

Talluri, B.C., Urai, A.E., Bronfman, Z.Z., Brezis, N., Tsetsos, K., Usher, M., Donner, T.H., 2020. Choices Change the Temporal Weighting of Decision Evidence. bioRxiv 1–16. doi:10.1101/2020.03.06.979690

van den Berg, R., Anandalingam, K., Zylberberg, A., Kiani, R., Shadlen, M.N., Wolpert, D.M., 2016a. A common mechanism underlies changes of mind about decisions and confidence. eLife 5, e12192.

Voss, J.L., Gonsalves, B.D., Federmeier, K.D., Tranel, D., Cohen, N.J., 2010. Hippocampal brain-network coordination during volitional exploratory behavior enhances learning. Nat Neurosci 1–9. doi:10.1038/nn.2693

Voss, M., Chambon, V., Wenke, D., Kühn, S., Haggard, P., 2017. In and out of control: brain mechanisms linking fluency of action selection to self-agency in patients with schizophrenia. Brain 140, 2226–2239.

Weiss, A., Chambon, V., Lee, J.K., Drugowitsch, J., Wyart, V., 2019. Interacting with volatile environments stabilizes hidden-state inference and its brain signatures. bioRxiv 4, 08825–56. doi:10.1101/755223

Wilson, R.C., Collins, A.G., 2019. Ten simple rules for the computational modeling of behavioral data. eLife 8, e49547.

Wilson, R.C., Geana, A., White, J.M., Ludvig, E.A., Cohen, J.D., 2014. Humans use directed and random exploration to solve the explore–exploit dilemma. Journal of Experimental Psychology: General 143, 2074–2081. doi:10.1037/a0038199

Wyart, V., Koechlin, E., 2016. ScienceDirect Choice variability and suboptimality in uncertain environments. Current Opinion in Behavioral Sciences 11, 109–115. doi:10.1016/j.cobeha.2016.07.003

Xu, F., Tenenbaum, J.B., 2007. Sensitivity to sampling in Bayesian word learning. Developmental Science 10, 288–297. doi:10.1111/j.1467-7687.2007.00590.x

Zylberberg, A., Wolpert, D.M., Shadlen, M.N., 2018. Counterfactual reasoning underlies the learning of priors in decision making. Neuron.

Acknowledgments

MR is the beneficiary of a postdoctoral fellowship from the AXA Research Fund. MR work was also supported by a department-wide grant from the Agence Nationale de la Recherche (ANR-17-EURE-0017, EUR FrontCog). This work has received support under the program «Investissements d'Avenir» launched by the French Government and implemented by ANR (ANR-10-IDEX-0001-02 PSL). A.W. was supported by the FIRE Doctoral School. V.C. was supported by the French National Research Agency (ANR-16-CE37-0012-01). J.D. was supported by the James S. McDonnell Foundation (grant #220020462). This work was supported by a starting grant from the European Research Council (ERC-StG-759341) awarded to V.W., a junior researcher grant from the French National Research Agency (ANR-14-CE13-0028-01) awarded to V.W., a France-US collaborative research grant from the French National Research Agency (ANR-17-NEUC-0001-02) and the National Institute of Mental Health (1R01MH115554-01) awarded to V.W. and J.D.

Author contributions

Conceptualization: V.W.; Data Collection: A.W., V.C., J.K.L., J.B., Methodology: A.W., V.C., J.D., and V.W.; Formal Analysis: M.R., A.W. and V.W.; Investigation: A.W., J.K.L., V.C., M.R. and V.W.; Writing – Original Draft: M.R. and V.W.; Writing – Review & Editing: M.R., V.C. and V.W.; Supervision: V.W.; Funding Acquisition: J.D. and V.W.

Conflicts of interest

The authors declare no competing interests.